

TOTAL NUMBER OF SEQUENCES

SEQUENCE LISTING - WEE virus strain 71V-1658 - SEQ ID NO 1

LOCUS AF214040 11484 bp RNA VRL 27-JAN-2000
DEFINITION Western equine encephalomyelitis virus strain 71V-1658 nonstructural polyprotein and structural polyprotein genes, complete cds.
ACCESSION AF214040 AF143811
VERSION AF214040.1 GI:6760410
KEYWORDS
SOURCE Western equine encephalomyelitis virus.
ORGANISM Western equine encephalomyelitis virus Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus.
REFERENCE 1 (bases 1 to 11484)
AUTHORS Netolitzky,D.J., Schmaltz,F.L., Parker,M.D., Rayner,G.A., Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.
TITLE Complete genomic RNA sequence of western equine encephalitis virus and expression of the structural genes
JOURNAL J. Gen. Virol. 81 Pt 1, 151-159 (2000)
MEDLINE 20109043
REFERENCE 2 (bases 1 to 11484)
AUTHORS Netolitzky,D.J., Schmaltz,F.L., Parker,M.D., Rayner,G.A., Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-1999) Medical Countermeasures Section, Defence Research Establishment Suffield, P.O. Box 4000, Stn Main, Medicine Hat, Alberta T1A 8K6, Canada
FEATURES Location/Qualifiers
source 1..11484
/organism="Western equine encephalomyelitis virus"
/strain="71V-1658"
/db_xref="taxon:11039"
5'UTR <1..24
misc_feature 1..11484
/note="sequence of this entry added to the sequence of AF214039, strain CBA87, constitutes the Western equine encephalomyelitis virus complete genome"

CDS 25..7428
/note="encodes nsP1, nsP2, nsP3 and nsP4 proteins"
/codon_start=1
/transl_except=(pos:5581..5583,aa:OTHER)
/product="nonstructural polyprotein"
/protein_id="AF28339.1"
/db_xref="GI:6760411"

/translation="MERI HVLDADSPYVKSLQRTFPQFEIEARQVTNDNHANARAFS
HVATK LIESEVDRDQVILDIGSAPVRHAHSNHYHCICPMISAEDPDRLQRYAERLKK

SDITDKNIASKAADLLEVMSPTPDAETPSLCMHTDATCRYFGSVAVYQDVYAVHAPTSI
YHQALKGVRTIYWIGFDTTPFMYKNMAGSYPTYNTNWADERVLEARNIGLGNSDLQES
RLGKLSILRKKRLQPTNKIIIFSVGSTIYTEDRSLLRSWHLPNVFHLKGKSNTGRCGT
IVSCEGYVIKKITISPGLYGVENLASTMHREGFLSKVTDTLRGERVSFAVCTYVPA
TLCQMTGILATDVSVDAAQKLLVGLNQRIVVNNGRTQRNTNTMQNYLLPVVAQAFSRW
AREHRADLDDEKELGVRERTLTMGCCWAFKTQKITSIYKKPGTQTIKKVPAVFDSFVI
PRLTSHGLDMGFRRLKLLEPTVKPAPAITMADVEHLRLQQEAEEVAAAEEIREAL
PPLLPEIEKETVEAEVDLIMQEAGAGSVETPRGHIRVTSPGEEKIGSYAILSPQAVL
NSEKLACIHPLAEQVLVMTHKGAGRGRYKVEPYHGKVIVPEGTAVPVQDFQALSESATI
VFNEREFVNRYLHHIAINGGALNTDEEYYKTVKTQDTDSEYVFDIDARKCVKREDAGP
LCLTGDLVDPFHFEPAYESLKTRPAAPHKVPTIGVYVGPGSGKSGTIKSAVTKKDLVV
SAKKENCAEIIRDVRRMRRMDVAARTVDSVLLNGVKHPVNTLYIDEAFACHAGTLLAL
IAIVPKKKVVLCGDPKQCGFFNNMCLKVHFNHDICTEVYHKSISRRCTQTVTAIVSTL
FYDKRMKTVNPCADKIIDTTGTTKPHKDDLILTCFRGVWKQLQIDYKHIMTAAAS
QGLTRKGVYAVRYKVNENPLYSQTSEHVNVLLTRTEKRIIWKTLAGDPWIKTLTAKYP
GDFTASLDDWQREHDAIMARVLDPQTADEFQNKVNVCWAKALEPVLATANIVLTRQQ
WE TLH PFKH DRAY SPE MAL NFF CTR FF GV DLS GL FSA PT VAL TY RD QH WD NS PG K NM
Y GLN REV AKEL S R R Y PC IT K A V D T G R V A D I R N N T I K D Y S P T I N V V P L N R R L P H S L I V D
HKGQGTTDHSGFLSKMKGSKVLVIGDPISTPGKKVESMGPLPTNTIRCDL LGIP SHV
GKYDIIFVNVRTPYRNHHYQQCEDHAIHHSMLTCKAVHHLNTGGTCVAIGYGLADRAT
ENIITAVARSFRFTRVCQPKNTAENTEVLFVFFGDNGNMHTDQDRLGVVLDNIYQGS
TRYEAGRAPAYRVIRGDISKSADQAIVNAAANSKGQPGSGVCGALYRKWPAAFDRQPIA
VGTARLVKHEPLIIHAVGPNSKMPPEPEGDLKLAAAYMSIASIVNAERITKISVPLLS
TGIYSGGKDRVMQSLHHLFTADFDTTADVTIYCLDKQWETRIIEAIHRKBSVEILDDD
KPVVIDLVRVHPNSSLAGRPGYSVNEGKLYSYLEGTRFHQTAKDIAEIHAMWPNKSEA

T007221-5419-2001

NEQICLYILGESMSSIRSKCPVEESEASAPPHTLPCLCNYAMTAERVYRLSAKKEQF
AVCSSFLLPKYRITGVQKLQCSKPVLPSGVVPPAVHPRKYAEIIILETPPPPATTVIC
EPTVPERIPSPVISRAPSAESLLSLGGVSFSSSATRSSTAWSDYDRRFVVTADVHQAN
TSTWSIPSAPGLDVQLPSDVTDSHWSIPSASGFEVRTPSVQDLTAECAKPRLAEMQ
DFNTAPFQFLSDYRPVPAPRRRPIPSPRSTASAPPVKPRTKYQQPPGVARAISEAE
LDEYIIRQHSNSRXYEAGAYIFSSETGQGHLOQKSVRQCKLQEPILDRAVHEKYYAPRLD
LEREKMLQKKLQLCASEGNRSRYQSRKVENMKAITAERLISGLGTYSSEVNVPVECYR
VNYPVPIYSSTVINRFTSAAEVAKTCNLVIQENYPTVASYCITDEYDAYLDMVGASC
CLDTATFCPAKLRSYPKHSYLQPEIRSAVPSPIQNTLQNVLAAATKRNCNVTQMREL
PVLDAAFNVDCFKKYACNDEYWDTFRDNPIRLTENVTVQYVTKLKGPKAAALFANTH
NLKPLQEIPMQDFVMDLKRDVKVTPGTTKHTERPKVQVIIQAADPLATAYLCGIHRELV
RRLNAVLLPNIHTLFDMAEDFDAIIAEHFHHGDPVLETDIASFDKSEDDAIAlAISLM
ILEDLGVDQPLLDLIEAAFGNTSVHLPTGTRFKFGAMMKSGMFLTFVNNTLVNIMIA
SRVLRERLTTSSACAASIGDDNIVHGVVSDTLMAERCATWLNMEVKIIDAVIGIKAPYF
CGGFILVDQITGTACRVADPLKRLFKLGKPLPVDDTQCDERRRALHDEAMRWNRIGIT
DELVKAVESRYEITILAGLIITSLSTLAESVKNFKSIRGSPITLYG"

CDS 7473..11183
/note="encodes nucleocapsid, E3, E2, 6K and E1 proteins"
/codon_start=1
/product="structural polyprotein"
/protein_id="AAF28340.1"
/db_xref="GI:6760412"

/translation="MFPYPQLNFPPVYPTNPMAYRDPNPPRRWRPFRRPLAAQIEDL
RRSIVNLTTFKQRSPNPPPGPPPKKKSAPPKPTQPKKKKQQAKRTKRKPCKPGKRQRM
CMKLESDKTFPIMLNGQVNGYACVVGGRLMKPLHVEGKIDNEQLAAVKLKKASMYDLE
YGDVPQNMKSDTLQYTSDKPPGFYNWHGAVQYENGRFTVPRGVGGKGDSGRPILDNR
GRVVAIVLGGANEGRTRTALSVVTWNQKGVТИRDTPEGSEPWSLVTALCVLSNVTFFCD

T01221.614292001

KPPVCYSLTPERTLDVLEENVDPNYDTLLENVLKPSRRPKRSITDDFTLTSPYLGF
CPYCRHSTPCFSPIKIENVWDSEDDGSIRIQVSAQFGYNQAGTADVTKFRYMSFDHDH
DIKEDSMEKIAISTSGPCRLGHKGYFLLAQCPPGDSVTVSITSGASENSCTVEKKIR
RKFVGREEYLFPVVHGKLVKCHVYDHLKETSAGYITMHRPGPHAYKSYLEEASGEVYI
KPPSGKNVTYECKGDYSTGIIVSTRTKMNMGCTAKQCIAYKSDQTWKVFNSPDILRHT
DHHSVQGKLHIFPRLPTVCPVPLAHTPTVKWFKGITLHLTAMRPTLLTRKLGLRAD
ATAEWITGSTSRNFSVGREGLEYVWGNHEPVRVWAQESAPGDPHGPHEIIIHYYHRK
PVYTIVLCGVALAILVGTASSAACIACKARDCLTPYALAPNATVPTALAVLCCIRPT
NAETFGETLNHLWFNNQFFLWAQLCIPLAALVILFRCFSCCMPLVAGVCLGKVDAF
EHATTVPNVPGIPYKALVERAGYAPLNLEITVVSSELTPSTNKEYTCKFHTVIPSPQ
VKCCGSLECKASSKADYTCRVPGGVYPFMWGAQCFCDSENTQLSEAYVEFAPDCTID
HAVALKVHTAAALKVGLRIVYGNNTAHLDTFVNGVTPGSSRDLKVIAGPISAAFSPFDH
KVVIRKGLVNYDFPEYGAMKPGAFGDIQASSLDATDIVARTDIRLLKPSVKNIHVPY
TQAVSGYEMWKNNNSGRPLQETAPFGCKIEVEPLRASNCAHYHIPISIDIPDAAFVRSS
ESPTILEVSTVADCIIYSADFGGSLTLQYKADREGHCPVSHSTTAVLKEATTHVTAV
GSITLHFSTSSPQANFIVSLCGKKSTCNAECKPPADHIIGEPHKVDQEFQAAVSKTSW
NWLLALFGGASSLIVVGLIVLVCSSMLINTRR"

3' UTR

11184..11484

BASE COUNT 3251 a 2842 c 2808 g 2583 t
ORIGIN

1 accttacaaa ctaatcgatc caatatggaa agaattcagc ttgacttaga tgctgacagc
61 ccgtatgtca agtcgttaca gggcacgtt ccacaatttg agatcgaagc aaggcaggtc
121 actgacaatg accatgccaa tgccagagcg ttttcgcattg tggcaacaaa gtcatttgag
181 agcgaagtgc accggggcca agttatcttgc gacattggaa gtgcgcctgt cagacatgca
241 catttcaatc accgcgtatca ttgttatctgc cctatgataa ggcgtgaaga cccggacaga
301 ctacaacgggt atcggaaaag acttaagaaaa agtgcatttgc acaccacgcg cagagactcc atctctgtgt
361 aaggcggcag actgtgttgc agtcatgtca
421 atgcacacag acggcacgtg tagttactttt ggaagtgttag cagtatccca agatgtgtac
481 gcagttccatg caccgcacatc aatcttacac caggcgctta aaggagttag gacaattac
541 ttgataggct ttgacacgc ccctttatg tacaaaaaca tggcagggttc ctaccctact

TOP SECRET

601	tacaacacga	actgggtga	cgagagaga	ttgaaagcac	gtaacattgg	cctcggtaac
661	tcagatcttc	aggagagcg	gttggaaaa	ctctcaatcc	ttaggaagaa	gaggctcaa
721	cctactaata	agatcatatt	ctcgggttgt	tcaacaatct	acacagaaga	tagatcaactg
781	ttacgtact	ggcatcttcc	aaacgtgttc	cacttgaagaa	gaaagtctaa	cttcacaggt
841	agatgtggaa	ccattgttag	ctgtgaaggg	taacgtcatca	aaaagataac	gatcaggcca
901	ggactatacg	gtaaaggta	gaacttggcg	tccacaatgc	atcgcgaggg	tttcttgagt
961	tgc当地tca	cagatacgct	gcccggcggag	agggtttctt	ttgctgttg	taegtatgt
1021	ccagccacac	tttgcgatca	gatgacagggg	attctggca	ctgacgttgc	tgtggatgac
1081	gcacaaaac	tattgggtgg	gtctcaacaa	aggattgtcg	tcaatggtag	gacgcaaaaga
1141	aatactaaca	caatcgacaa	ctatcttata	ccagtgtcg	ccccaggggt	ttccagggtg
1201	gcccgtgaac	atctgtttcc	cttggacgac	gagaagaaac	taggggtcg	ggagcgcact
1261	cttactatgg	gtctgtctgt	ggcttcaag	accggaaaaaa	tcacatccat	ctacaagaag
1321	cctggtaacgc	aaacaattaa	gaaaggtaact	gcccgttgg	actcattttgt	gattccacgc
1381	cttaccaggcc	acgggtttcc	tatgggttc	cgccgttaggc	tcagactgt	gcttgc当地a
1441	actgtcaaac	ccgcacccggc	tattacaatg	ccgcgttgcg	agacatcgcg	tggcttacag
1501	caagaagctg	aagaagtggc	tgcagcgaa	gagatcgac	aaggccgtcc	acecttgc
1561	cctgaaatag	aaaaagagac	cgtagaggac	gaagttagac	tcattatgc	agaggcaggaa
1621	gcaggttagc	tggagacacc	acggagacac	atcgggtga	caagtttacc	aggcgaagag
1681	aaggatgggt	cttgcgtat	attttacccc	caggcgttat	tgaatagtga	aaaacttggcg
1741	tgtatccacc	cattggcgga	acaaggtaact	gtaatgact	acaaaaggtag	ggcaggggaga
1801	tacaaggatcg	acccataccac	cggtaaaggc	attgttaccc	aaggggacgc	gggtccctgtt
1861	caagaaatcc	aggccatggc	tgagagcgct	acgatctttt	tcaacgagac	ggagtttgc
1921	aacagatacc	tgccaccat	cgcaatcaac	ggaggagcgc	taaacactga	cgaagagatc
1981	tataaqtact	taaagactca	ggacacagac	tcgaaatact	tcttgcata	tgacgc当地a
2041	aagtgttta	acggagaaaga	cgccaggctcc	ttgtgc当地	ccgggttact	ggtatgatca
2101	ccatttcacg	agtttgcgt	cgagagtctc	aagacacgc	cagcagcacc	tcacaaaagtc
2161	ccaaaccatcg	gagttctatgg	agtgcggatgt	tcgatgttat	ctggaaatcat	ccaaaggct
2221	gtgactaaga	aagatctgtt	tgtgtgtcg	aagaaggaaa	atcgcgc当地	aatcatcagg
2281	gatgtaaggaa	ggatgagacg	tatggatgtt	gtctgttagga	ctgtcgattc	agtgttctt
2341	aatggggtta	acgaccccg	taacactctg	tacattgtat	aggcatttgc	ctggccatgca
2401	gggaacgtgc	tggcactgt	tgccatgtc	aaacctaaga	aagtgttgc	tgccggggac
2461	ccaaaacaat	ggccgttctt	taacatgtat	tgccgttgc	tgccgttgc	ccatgacata
2521	tgcaactgaag	tgtaccataa	aagcatcttct	aggaggtgc	cacagactgt	aaccgc当地
2581	gttccacgc	tcttcgtca	caagcgatgt	aagacgttgc	accgcgttgc	tgataaaatc
2641	atcatagata	ccacaggccg	cacaaaaggcc	cacaaaatgt	atctgttctt	aacctgttcc
2701	agaggatggg	tggaaacagct	acagattgac	tacaaaatc	acgaaatcat	gactgc当地
2761	gcatcgcaac	gacttacgcg	gaaagggttgc	tatgtgttgc	ggtacaaaat	caacagaaat
2821	ccacttctact	cgccagactt	tgagacgtgt	aacgttgc	ttacacggcc	agaaaaaacgc
2881	attgtcttgg	agacgttgc	tggtgtatccc	tggataaaga	cacttacagc	taaatatccc
2941	ggggatttca	cggttcttgc	ggacgtactgg	cagcgc当地	acgacgc当地	tatggc当地
3001	gttcttgcata	aggcgccagac	agctgtatgt	tcccaata	agggttgc	ctgttggcg
3061	aagggtttag	agccgttctt	ggccacggcc	aacattgtgc	tgacgagaca	gcagtgggag
3121	acgttgcacc	cattcaagca	tgacagacggc	tactcactt	aatggcact	gaacttctt
3181	tgc当地gg	tcttgcgtt	agacgttggac	agtgggttat	tttccgttcc	tacgttgc当地
3241	cttacttatac	gggatcagca	ctgggataac	tcgc当地gg	agaacatgt	tgggttcaat
3301	agagaggtag	caaaggatgt	gtc当地gg	tatccgtca	tcacaaaaggc	ggttgc当地
3361	ggccagggtt	ctgtatataa	gaaataacc	atcaaggat	actcttccaa	aataatgt
3421	gttccattaa	atccggccgtt	gccccactcg	ttgtgttgc	accacaaaagg	acagggttca
3481	actgatcaca	ggggatttctt	atctaagatg	aaggc当地	ctgtgttgc	gatcggc当地
3541	cctatcagca	ttccaggggaa	gaaagttagag	tccatgggtc	cattggccac	taataaccatc

6601 ctgaatgcgg tgcttctgcc aaatatccat actctcttcg acatgtcagc ggaagatttt
 6661 gatgcgatta ttgctgaaca tttccaccac ggcgaccagg tattggaaac ggacatcgcg
 6721 tcgtttgata aaagcgaaga cgacgctatc gccatttcgg cggtatgtat ccttgaggac
 6781 tttagtgtcg accaacecgct tttagatttgc atagaggccgg cgttcggeaa tttcacatct
 6841 gtgcacctac ctacaggaaac gaggtttaaa ttgggtgcga tgatgaaatc cggtatgttc
 6901 ttaacgctgt ttgtcaacac actagtcaat atcatgatttgc atagcagatc actacgtgaa
 6961 cggtaacca cgtcagcgtg cggccgcctt atcggcgacg ataacatagt gcattgtgc
 7021 gtctccgaca ccttgcgtgc gggagatgc gccacttggc tgaacatggaa agaaaaatt
 7081 attgtgcag ttattgttat caaaggcccc tacttctgtt ggggattttt cttggggc
 7141 cagataaacag gcacagcgtg cagagtgcga gaccctttaa aaaggctttt taagcttgg
 7201 aaaccattgc cgtcgtatc taccacaaacatc tgccggcact gecgggcact gcatgtgaa
 7261 gcaatcgcatc ggaacacaat tggattttacg cggacttgc tgaaggccgt agaatccaga
 7321 tacgagatca tactggcagg octgtatcact acgtctgtt ccacgttgc cggaaacggtt
 7381 aagaacttca agacatcaat aggaggccca atcccccattt acgggttgc gaaatgg
 7441 acgttagtaa caccgcattttt cccacccggc gaatgtttc atacccttgc ctgaaacttc
 7501 caccagttt cccttacaaat ccgtatggctt acccgatc aaacccttctt aggcgcgc
 7561 ggaggccgtt tcggccccccg ctgggtgtc tcaatcataatc ttttaggggg ttttttttt
 7621 actgtacttca caacaacatc tcaatcataatc cggccgcctt tccacccggca aagaagaaga
 7681 agagtgtctcc taaggccaaaat cttacttcgc tttttttttt tttttttttt
 7741 cggaaacgcaat gcttaaccca gggaaacgac aacgtatgtt tatgtatgtt gatgtatgtt
 7801 agacattttc gatctgtcg aacggccaaat tttttttttt tttttttttt
 7861 ggctgtatggaa accactccac gttgtggaa aattgtatggaa ttgtatggaa
 7921 aattgtatggaa ggtagatcatc tttttttttt tttttttttt
 7981 cagacacgtc ggtatgcacc acggacaaatc tttttttttt tttttttttt
 8041 cttttttttt tttttttttt tttttttttt tttttttttt
 8101 gcgaaagacc gatctggac aacagaggca tttttttttt tttttttttt
 8161 atggggccac ggtagccggc tttttttttt tttttttttt
 8221 gggatcccccc cggaaatgttcc gaaatgttcc tttttttttt tttttttttt
 8281 atgtacatgtt cccatgtgcac aaccccccggc tttttttttt tttttttttt
 8341 tcgactgtgtt cggaaatgttcc gttttttttt tttttttttt
 8401 tggaaatgttcc atacgcggccgg cccaaatggaa tttttttttt tttttttttt
 8461 ctttacttggg gtttctggcc tttttttttt tttttttttt
 8521 ttgagaacgtt gttttttttt tttttttttt tttttttttt
 8581 tcggctatcatc tcggccggc tttttttttt tttttttttt
 8641 acgacatcatc catcaaggaa gacatgtatgg tttttttttt tttttttttt
 8701 gccgtcgatc tggccacaaa gggatgttcc tttttttttt tttttttttt
 8761 taacgttccatc tttttttttt tttttttttt tttttttttt
 8821 ggaggaaatgttcc gggatgttcc tttttttttt tttttttttt
 8881 atgtccacgtt tttttttttt tttttttttt tttttttttt
 8941 cggccccaca cggatgttcc tttttttttt tttttttttt
 9001 caccttctgg caagaacgtc acatgtatgg tttttttttt tttttttttt
 9061 tgacacgtcg aacaaatgttcc tttttttttt tttttttttt
 9121 gcgacccaaatc gaaatgggttcc tttttttttt tttttttttt
 9181 tggaaatgttcc atttgtatgttcc tttttttttt tttttttttt
 9241 cttttttttt tttttttttt tttttttttt tttttttttt
 9301 gaccaacatgttcc gttttttttt tttttttttt tttttttttt
 9361 ttacagggttcc tttttttttt tttttttttt tttttttttt
 9421 gtaacatgttcc accatgtatgttcc tttttttttt tttttttttt
 9481 ggccgtatgttcc gttttttttt tttttttttt tttttttttt
 9541 tttttttttt tttttttttt tttttttttt tttttttttt

9601 aagaagaag agactgcctg acgccatacg cgcttgaccc gaacgcaacg gtacccacag
9661 cattagcgg tttgtgtgc attcggccaa ccaacgctga aacatttga gaaacttta
9721 accatctgtg gttAACAC caacggttc tctgggcaca gttgtgcatt cctctggcag
9781 cgctttagt tctgttccgc tgctttcat gtcgcattgc ttttttatgg ttgcaggcg
9841 tctgcctggg gaaggtagac gccttcgaac atgcgaccac tgcgccttgc aatccaccc
9901 tcccgataa ggccgttggc gaacgcgcag gttacgcgc acttaaccc gagatcacgg
9961 tgcgttcatac ggaattaaca cttcaacta acaaggagta cgtgacccgc aaattccaca
10021 cagtcattcc ttacccacaat gttaaaatgtc gccgggtccct cgagtgcacag gcatcccaa
10081 aggggattt cacatgccgc gtttttggcg gtgtgtaccc tttcatgtgg ggagggcgc
10141 aatgcgttctg tgacagttag aacacacaat tgatgtggc gtagtgcggg ttgcgttccag
10201 atgcgtactat agatcacgcg gtcgcactaa aagttcacac agtcgtctg aaatgcggcc
10261 tgcgtatagt atacggcaac accaccgcgc accttgcatac gtttgtcaat ggccgtcacgc
10321 cagggttccctt acgggacccgt aagggtcatag cagggccgcg atcagccgc ttttcaccc
10381 ttgaccataa ggctgttcatac agaaaggggc ttgttttacaa ctacgactc cctggatgt
10441 gagctatgaa acccaggagcg ttccggcgata ttcaagcatac ctcgcgtgt gtcacagaca
10501 tagtagcccc cactgacata cggctgtca agccttctgt caagaacate cacgtccct
10561 acacccaacg agtatccaggg tatgaaatgt ggaagaacaa ctcaggacga cccctgcgaag
10621 aaaaacgacc accatggatgt aaaaatttgaag tggagccctt ggcagcgtct aactgtgttt
10681 acgggcacat ccctatctcg attgacatcc ctgatgcacgc ttttgtgaga tcatacgaat
10741 caccacaaat tttagaaatgt agctgcacag tagcagactg catttatttc gcagacttgc
10801 tggtttctt aacattacag tacaagatgt acaggggggg acatgttccca gttcaactccc
10861 atccacacg agctgttttga aagaaaggcgca ccacacatgt gactgcgtt ggcagcataa
10921 cactacatcc tagcacatcg agccccaacaa caaatttat agtttcgccta tgccggcaaga
10981 agtcacccgt caatgtgtaa tgtaaaccac cggccgcacca cataatttggaa gaaccacata
11041 aagtccacca aagaattcccg gccggcgtttt caaaaacacatc ttggaaatggg etgcgtgcac
11101 tgggggggg agacatcatcc ctcatgtttg taggacttat agtgttggc tgcaagctcta
11161 tgcttataaa cacacgtata tgacttgaggc cgggactgtg cataggggtt aaactcgatg
11221 tactcccgag gaagcgttggt gcataatggc acgcgcgcgt tgacactaaa actcgatgt
11281 ttcccgagga agacacagtgc ataatgtgt gcaagtgcac attaattcgta tatacactac
11341 catattaaaca acactatatac acctttatgt gactcaactat gggctcttaa tatacactac
11401 acatattttt cttaaaaaca ctatacacac ttatataattt cttttataat ttttttttt
11461 tttttttttt gttttttttt ttcc
//

SEQUENCE LISTING - pcdWXH-7 - SEQ ID NO 2

<210> SEQ ID NO 2
<211> LENGTH: 4150
<212> TYPE: DNA
<213> ORGANISM: WESTERN EQUINE ENCEPHALITIS VIRUS
<214> FEATURES:
Location/Qualifiers
source 1..4150
/organism="Western equine encephalomyelitis virus"
/strain="71V-1658"
vector sequence 1-9
5' SacI primer 9-20
3' end of NS4 gene 16-114
intragenic region 115-159
CDS 160-3869
/note="encodes capsid, E3, E2, 6K and E1 proteins"
/codon_start=160
/product="structural polyprotein"

/translation="MFYPQLNFPVYPTNPMAQRDPNPPRRWRPFRRPLAAQIEDL
RRSIVNLTFKQRSPNPPGPPPKKKSAPKPKPTQPKKKQQAKRTKRKPCKPGKRQRM
CMKLESDKTFPIMLNGQVNGYACVVGGRLLMKPLHVEGKIDNEQLAAVKLKKASMYDLE
YGDVPQNMSKDTLQYTSDKPPGFYNWHGAVQYENGRTVPRGVGGKGDSGRPILDNR
GRVVAIVLGGANEGRTRTALSVTWNQKGVTIRDTPEGSEPWSLVTALCVLSNVTFPCD
KPPVCYSLTPERTLDVLEENVDNPNEYDTLLENVLKCPSPRKRSITDDFTLTSPYLGF
CPYCRHSTPCFSPIKIENVWDESDDGSIQVSAQFGYNQAGTADVTKFYRMSFDHDH
DIKEDSMSKEIAISTSGPCRRLGHKGYFLLAQCPCPGDSVTVSITSGASENSCTVEKKIR
RKFVGREEYLFFFVHGKLVKCHVYDHLKETSAGYITMHRPGPHAYKSYLEEASGEVYI
KPPSGKNVTYECKCGDYSTGIVSTRTKMNGCTAKQCIAYKSDQTWKWFNSPDLRHT
DHSVQGKLHIPFRLPTVCPVPLAHTPTVTKWFKGITLHLTAMRPTLLTRKLGLRAD
ATAEWITGSTSRNFSGREGLEYVVGHNHEPVRVWAQESAPGDPHGPHEIIIHYYHRH
PVYTVIVLCGVALAILVGTASSAACIKAARRDCLTPYALAPNATVPTALAVLCCIRP
NAETFGETLNHLWFNNQPFLWAQLCIPLAALVILFRCFSCCMPFLLVAGVCLGKVDAF
EHATTVPNVPGIPYKALVERAGYAPLNLEITVVSSELTPSTNKEYVTCKFHTVIPSPQ

VKCCGSLECKASSKADYCRVFGGVYPFMWGAQCFCDSENTQLSEAYVEFAPDCTID
 HAVALKVHTAALKVGLRIVYGNTTAHLDTFVNGVTPGSSRDLKVIAGPISAAFSFHDH
 KVVIRKGVLVNYDFPEYGMKPGAFGDIQASSLDATDIVARTDIRLLKPSVKNIHVYPY
 TQAVSGYEMWKNNSGRPLQETAPFGCKIEVEPLRASNCAYGHIPISIDIPDAAFVRSS
 ESPTILEVSTVACDIYSADFGSSTLQYKADREGHCPVSHSTTAVLKEATHVTAV
 GSITLHFSTSSPQANFIVSLCGKKSTCNAECKPPADHIIGEPHKVDQEFOAAVSKTSW
 NWLLALFGGASSLIVVGLIVLVCSSMLINTRR"

3' UTR

3870-4150

BASE	COUNT	ORIGIN
1	ggcccttag	agctcatact
61	agcgtaaga	acttcagaag
121	tagtgacgt	agttagacacg
181	actttccac	agtttacccat
241	gccgcgttag	ggcgttccgg
301	tagtcaactt	gactttcaaa
361	agaagaagag	tgtcttaaag
421	aggaggcga	acgcgaacgt
481	cggacaagac	atttccgatc
541	gaggaaaggct	gatgaaacc
601	ccgtgaaattt	gaagaaggct
661	tgaatcaga	cacgctgcag
721	acggcgcagt	ccgatcttag
781	gcgcacgcgg	aagacgcatt
841	gtgcataatg	gggcacgcgt
901	ccatctaggaa	taccccccgg
961	tttcaatgt	cacgttccca
1021	gaacactcga	cgtgtccga
1081	acgttctgaa	atgtccatca
1141	ccagtccctt	cctgggggtt
1201	aaaaatttga	aaacgttgg
1261	cacaattcgg	ctacaatcag
1321	tcgaccacgg	ccatgacatc
1381	gaccctgcg	tcgttggc
1441	acagtgtaac	cgtcgtatc
1501	agatcaggag	gaagtggcgtt
1561	tggtaaatgt	ccacgtttac
1621	acaggccagg	cccacacgcg
1681	ttaaacacc	ttctggcaag
1741	gtatctgt	cacgcgaacg
1801	aaacagcga	ccaaacggaa
1861	actcaatgt	ttgggtttca
1921	cgttagctca	cacgcttaca
		gtcacgaagt
		ggttcaaagg
		catccccctc
		cacctgactg

1981 caatgcgacc aacattgtctg acaacggaaa aattggggct gcgagcagac gcaacagcag
 2041 aatggattac aggttctaca tcaggaaatt ttctgtggg gcgagaagg ctggagtacg
 2101 tatggggtaa ccatgaacca gtcagagttt gggcccagga gtcggcacca ggcgaccac
 2161 atggatggcc gcatggatc atcatccact attatcatcg gcatccagtc tacactgtca
 2221 ttgtgtctgt tggtgtcgct ctgtctatcc tggttaggcac tgcatcatca gcacgttgca
 2281 tgcgccaagg aagaagagac tgctgtacgc catacgctgct tgacccgaaac gcaacggatc
 2341 ccacaggattt aacacccatc ggccaaacca cgctgttttgc acatccatc
 2401 ctggaaacca tctgtgggtt aacaaccaac cggttctctg ggcacagttt tgcatcttc
 2461 tggcagcgct tggatccatc ttccgtgtc ttccatgtcg catgcctttt ttatgttg
 2521 caggcgcttg ctggggaaag gttagacgcct tcgaaacatgc gaccactgtg ccaaattgttc
 2581 cggggatccc gtataaggcg ttgtgtcgac ggccgggtt cgccgactt aacccggaaa
 2641 tcacggctgt ctatcgaaatc ttacacccatc caactaaacaa ggatgttgtt acctgtcaaa
 2701 tccacacagt catccatc ccacaaggta aatgtgtcg gtcctcgat tgcaaggcat
 2761 ctcggaaaggc ggattacaca tggcgctgt tgccgtgt tgaccctttt atgtggggag
 2821 ggcacaaatg ctgtgtgac agtgagaaaca aacaaatcgat tgaggcgatc gtccgtgtcg
 2881 ctccagactg cactatagat caccgtcgct cactaaatgt tcacacatcg tgcgttggaaa
 2941 tcggcgtcg tatagtatc gcaacacca cccgcgcacct ggatacggtt gtcattgtcg
 3001 tcacggccagg ttcttcacgg gacctgtggg tcatcgagg gccgatatacg gccgttttt
 3061 cacccttgc ccataagggtc gtcatcgaaa aggggttgc ttacaactac gacttccctg
 3121 agtattggacg tatgaaacca ggacgttgc gcgatattca agcatctcg ttgtatgtca
 3181 cagacatcgat agcccgact gacatcgcc tcgtgtggaa ttctgtggaa aacatccacg
 3241 tccccatcac ccacggatc tcagggtatc aaatgtggaa gacaactca ggacgacccc
 3301 tgcaagaacac agcaccatggatgtaaaa ttgaagtggg gctctcgatc gctctact
 3361 gtgtttccgg gacatccctt atctcgatgg acatccctgtc tgccgtttt tgagatcat
 3421 cagaatcacc aacaattttta gaagttatgc gacacatgc agactgcatttattgtcg
 3481 actttgtgg ttctctaaca ttacagtaca aagctgacag ggaggggatc tgccgttgc
 3541 actccactc cacacatgtt gtttggaaagg aacgcgaccac acatgtgactt gccgttaggc
 3601 gcataaaactt acatccatgc acatcgaccc cacaaggaaa tttttagtgc tgctatgtcg
 3661 gcaagaatgc cactgtcaat gtcgtatgtc aaccacccggc cgaccacata attggagaac
 3721 cacataaagt cgaccaagaa ttccaggccggc cagtttccaa aacatcttgg aactggctgc
 3781 ttgcgtatgtt tggggggatc tcatccctca ttgtgttagg actttagtgc ttgggtctgca
 3841 gctctatgtc tataaaacaca cgtatgtac tgacgcggc cactgtacata gccgtaaaac
 3901 tcgtatgtact tccggaggaaag cgtgggtcat aatgccacgc gccgcttgac actaaaactc
 3961 gatgtatgtt cgagggaaagca cagtgcataa tgctgtcgat tgtcacatca atcgatatac
 4021 acactacata ttacaaacac tatatcatctt ttatgtgactt cactatgggt ctctataatc
 4081 cactacacat attttactta aaaacactat acacacttta taaattctt cataatttca
 4141 cttaggttt
 //
 48

SEQUENCE LISTING - pVHX-6 - SEQ ID NO 3

```
<200>      SEQUENCE CHARACTERISTICS
<210>      SEQ ID NO 3
<211>      LENGTH: 1260
<212>      TYPE: DNA
<213>      ORGANISM: EXAMPLE OF A CMV PROMOTER COUPLED TO THE pcDW-XH7
SEQUENCE
<220>      FEATURE:
  1-302          pVAX vector sequence
  302-401        CMV promoter
  411            CMV putative transcriptional start site
  444-463        T7 promoter
  464-509        pVAX multicloning region
  510            start of pcDW-XH4 sequence
```

1	cattgacgtc	aatgggtga	gtatttacgg	taaactgccc	acttggcagt	acatcaagtgc
61	tatccatatgc	caagtagcc	ccttattgtac	gtcaatgcg	gtaaatggcc	cgcttggcat
121	tatggccagg	atcatgcatt	atggggacttt	ctacttggc	agtatcatca	cgtatttagte
181	atcgcttata	ccatgggtat	gggggttttg	cagtagatca	atggggctgg	atagcggttt
241	gactcacggg	gatttccaag	tcccccacccc	attgacgtca	atggggaggtt	gttttggcac
301	caaatacAAC	ggggactttcc	aaaaatgtgt	aacaactcc	ccccatggc	gcaaaatggc
361	gggttaggcgt	tacgggtggga	ggttctata	acgagagtc	tctggtaac	tagagaaccc
421	actgttact	ggcttatacga	aaattaatcag	actcactata	gggagaccca	agctggctag
481	cggtttaaac	taaagcttggt	accggacttc	tactggcagg	cctgtatcatc	acgtctctgt
541	ccacgttagc	cgaaaagcgt	aagaacttc	agagactaaag	aggggaccca	atccacccct
601	acggctgacc	taataatggt	acgttagtaga	cacgcaccta	ccccacgc	aatgtttcc
661	atacccttag	ctgaaaccttc	caccaggfta	ccttacaaat	ccgatggott	acccgagatcc
721	aaacccttc	atggcgcgcgt	ggaggcccggt	tcggggccccc	ctggcgctgc	aaatcgaaga
781	tcttaggagg	tgcatagttc	acttgtaccc	caacacaacga	tcaccaatac	cgccggccgg
841	tccacccgcca	aagaagaaga	agagtgtcc	taaggccaaa	cctactcage	ctaaaaagaa
901	gaaggcaggaa	gccaaggaggaa	ggaaacggcaa	gcctaaacca	ggggaaacgc	aacgtatgt
961	tatggatgt	gagtccggaca	agacatttc	gtatcatgt	aacggccaa	tgaatggata
1021	tgccctgggt	gtcgaggaa	ggctgtatgg	accactccac	gttgaaggaa	aaattgtataa
1081	tgagcaat	ggccgcgtga	aatttgaaagaa	ggcttagcatg	tacgacttgg	agtaacggcga
1141	cgtttccccc	aactatggat	cgacacgcgt	cgatgtaccc	agcgcacaaa	caccgggggt
1201	ctacaaatgg	caccacggcc	cagtcccaata	tgagaatggg	aqatattac	taccqagagg

... CONTINUOUS WITH pcDW-HX45 (pcDW-XH7) SEQUENCE ...

4177 gaatgttaac cacccggccga ccacataatt ggagaaccac ataaagtca ccaagaattc
4237 caggcggcag ttcccaaaaac atcttggAAC tgctgtttg cactgtttg gggagactca
4297 tcccttcattt tgtaggact tatagtgttg tgctgcagct ctatgtttat aaacacaegt
4357 agatgtactga gcgcggacac tgacatagcg taaaactcg atgtacttc gaggaaaggt
4417 ggtgcataat gcacacgcgaa gtttgacact aaaactcgat gtatttccga ggaagcacaG
4477 tgcataatgc tggtgcagtgt cacattaatc gtatatacaca ctacatattttaaaacactat
4537 atcatctttt tgagacttcac tatgggtcte taatatacac tacacatattttttaatctttaa
4597 acactatatac cactttataa attttttttt aatttttttt tgctttttaqg qcacacttgc

4657 ggccgttact agtggatccg agctctagag ggcc

4177-4643 pcDW-HX45

4644-4790 vector sequence //

101011-6122101
10023649